

### **REMARKS**

In the non-final Office Action, the Examiner withdrew claims 12-15 from consideration due to a previous restriction requirement; indicated that two documents submitted with a previously filed Information Disclosure Statement (IDS) were missing from the Patent Office's file; indicated that Figures 1-3 should be designated with a "Prior Art" legend; rejected claims 18 and 19 under 35 U.S.C. § 112, second paragraph, as indefinite; rejected claims 1-3, 9, and 16 under 35 U.S.C. § 102(b) as anticipated by Colbath et al. ("Spoken Documents: Creating Searchable Archives from Continuous Audio," 2000) (which the Examiner identified as "Kubala et al."); rejected claims 4-8, 10, 11, and 17 under 35 U.S.C. § 103(a) as unpatentable over Colbath et al. in view of Liddy et al. (U.S. Patent No. 5,963,940); and rejected claims 18 and 19 under 35 U.S.C. § 102(b) as anticipated by, or in the alternative, under U.S.C. § 103(a) as unpatentable over Colbath et al.

By this Amendment, Applicants amend claims 18 and 19 to improve form. Applicants respectfully traverse the Examiner's request for a "Prior Art" legend and the Examiner's rejections under 35 U.S.C. §§ 112, 102, and 103. Claims 1-19 remain pending, of which claims 12-15 have been withdrawn from consideration by the Examiner.

### ***MISSING IDS DOCUMENTS***

In paragraph 3 of the Office Action, the Examiner indicated that two documents on a previously filed IDS were missing from the Patent Office's file. Applicants submit herewith copies of these documents. Accordingly, Applicants respectfully request that the Examiner consider these documents, initial at the appropriate place on the Form 1449, and return a copy of the Form 1449 with the next communication. A clean copy of the Form 1449 is attached hereto.

### *DRAWINGS*

In paragraph 5 of the Office Action, the Examiner indicated that Figures 1-3 should be designated by a "Prior Art" legend because only that which is old is allegedly illustrated. Applicants respectfully disagree.

Figures 1-3 are described in the Detailed Description section of the specification. As described in Applicant's specification, Figure 1 illustrates an exemplary system in which systems and methods consistent with the principles of the invention may be implemented (paragraph 0027). Applicants describe that the system of Figure 1 includes a server system, which is further described with regard to Figure 2. Applicants describe that the server system includes a server, which is further described with regard to Figure 3. These components are described in Applicants' specification as implementing certain functions consistent with the principles of the invention (see, e.g., paragraphs 0048-0060). Therefore, Applicants submit that Figures 1-3 do not constitute prior art.

Accordingly, Applicants respectfully request that the requirement for a "Prior Art" legend be reconsidered and withdrawn.

### *REJECTION UNDER 35 U.S.C. § 112, SECOND PARAGRAPH*

In paragraph 7 of the Office Action, the Examiner rejected claims 18 and 19 under 35 U.S.C. § 112, second paragraph, as indefinite for allegedly failing to particular point out and distinctly claim the subject matter that Applicants regard as the invention. Applicants respectfully traverse the rejection.

The Examiner alleged that "predetermined portion" recited in claim 18 lacks antecedent basis. Without acquiescing in the Examiner's rejection, Applicants have amended claim 18 to recite

"at least a predetermined number of the documents." Applicants submit that claim 18, as amended, is definite and satisfies the requirements of 35 U.S.C. § 112, second paragraph.

Accordingly, Applicants respectfully request the reconsideration and withdrawal of the rejection of claims 18 and 19 under 35 U.S.C. § 112, second paragraph.

*REJECTION UNDER 35 U.S.C. § 102(b)*

In paragraph 9 of the Office Action, the Examiner rejected claims 1-3, 9, and 16 under 35 U.S.C. § 102(b) as allegedly anticipated by Colbath et al. Applicants respectfully traverse the rejection.

A proper rejection under 35 U.S.C. § 102 requires that a single reference teach every aspect of the claimed invention. Any feature not directly taught must be inherently present. In other words, the identical invention must be shown in as complete detail as contained in the claim. See M.P.E.P. § 2131. Colbath et al. does not disclose or suggest the combination of features recited in claims 1-3, 9, and 16.

Independent claim 1, for example, is directed to a method of creating labels for clusters of documents. The method comprises identifying topics associated with the documents in the clusters; determining whether the topics are associated with at least half of the documents in the clusters; adding ones of the topics that are associated with at least half of the documents in the clusters to cluster lists; and forming labels for the clusters from the cluster lists.

Colbath et al. does not disclose the combination of features recited in claim 1. For example, Colbath et al. does not disclose clusters of documents. Colbath et al. mentions the word "cluster" in a few places in the context of clustering speakers together (see, e.g., column 5, lines 21-32), but does not disclose clusters of documents.

Because Colbath et al. does not disclose clusters of documents, Colbath et al. cannot disclose determining whether topics are associated with at least half of the documents in the clusters, as further recited in claim 1. The Examiner alleged that Colbath et al. discloses this feature and cited column 5, lines 34-48, and column 8, lines 11-16, of Colbath et al. for support (Office Action, page 4). Applicants disagree.

At column 5, lines 34-48, Colbath et al. discloses:

For the speaker change detection problem, we are able to find 90% of all speaker changes in general Broadcast News data within 100 milliseconds of the true (human-labeled) boundary. Accurate speaker identification is a much harder problem. The current system only has 20 known speakers (primarily well known anchor speakers and common news personalities) out of a speaker population of 170, but there are plans to increase this to about 100. With 20 known speakers, the system has a false rejection rate (that is, where audio segments of the target speaker are labeled anonymously) of 5%, and a false acceptance rate (where audio segments of the target speaker are labeled as another known speaker) of 2%.

In this section, Colbath et al. discloses identifying changes in speakers within an audio signal.

Nowhere in this section, or elsewhere, does Colbath et al. disclose determining whether identified topics are associated with at least half of the documents in the clusters, as required by claim 1.

At column 8, lines 11-16, Colbath et al. discloses:

The Rough'n'Ready IR system uses a full-text search system developed at BBN which uses an HMM-based model of document retrieval. This system, described in [7], is used in relevance-feedback mode to allow the user of the system to find documents that are similar to an exemplar.

In this section, Colbath et al. discloses an HMM-based model for document retrieval. Nowhere in this section, or elsewhere, does Colbath et al. disclose determining whether identified topics are associated with at least half of the documents in the clusters, as required by claim 1.

The Examiner also alleged that "a response is half-correct when the label is correct" (Office Action, page 4). Applicants cannot understand what the Examiner is alleging. If the Examiner

persists with a rejection of the above-identified claim feature based on Colbath et al., Applicants respectfully request that the Examiner provide a reasonable explanation why the Examiner believes that the above-identified sections of Colbath et al. disclose determining whether identified topics are associated with at least half of the documents in the clusters, as required by claim 1.

Colbath et al. also does not disclose forming labels for the clusters from cluster lists that include topics that are associated with at least half of the documents in the clusters, as further recited in claim 1. Instead, Colbath et al. discloses forming a title for a document from all of the topics associated with that document (column 10, lines 13-20).

The Examiner alleged that Colbath et al. discloses forming labels for the clusters from cluster lists that include topics that are associated with at least half of the documents in the clusters, and cited column 5, lines 21-32, column 8, lines 11-16, column 14, lines 29-33, and column 15, lines 1-10, of Colbath et al. for support (Office Action, page 4). Applicants disagree.

At column 5, lines 21-32, Colbath et al. discloses:

Speaker identification and segmentation allows us to create paragraph-like units between speakers, both known and unknown (classified by gender). Speaker change detection is important for correct playback of audio sections of the archive, and speaker identification allows the user of the archive to perform queries about particular speakers known to the system, and to skim over the archive for areas where particular speakers were present. In general, however, most speakers will be unknown to the system, so the speaker identification system will cluster them and give them a unique name.

In this section, Colbath et al. discloses speaker identification and segmentation that clusters unknown speakers and gives them a unique name (e.g., male 8 or female 6 - column 10, lines 22-27). Nowhere in this section, or elsewhere, does Colbath et al. disclose forming labels for the clusters from cluster lists that include topics that are associated with at least half of the documents in the clusters, as required by claim 1.

Column 8, lines 11-16, of Colbath et al. is reproduced above. In this section, Colbath et al. discloses an HMM-based model for document retrieval. Nowhere in this section, or elsewhere, does Colbath et al. disclose forming labels for the clusters from cluster lists that include topics that are associated with at least half of the documents in the clusters, as required by claim 1.

At column 14, lines 29 - column 15, line 10, Colbath et al. discloses:

Annotation

There is no particular reason that the database has to be browsed in a read-only fashion, however. The training data for the Rough'n'Ready indexer is currently fairly static. To annotate more speech data, or additional names for the name spotter, or additional topics for the topic classifier is a separate, offline process using dedicated annotators. However, since the current annotation process is relatively simple and does not require any in-depth linguistic knowledge, it seems logical that the enduser of the archive should be enlisted in helping to provide the training data. This makes sense since it is likely the consumer of the data will have the most familiar with it, and will be able to provide topics, identify speakers, etc.

In this section, Colbath et al. discloses that annotators can annotate more speech data, additional names for the name spotter, or additional topics for the topic classifier. Nowhere in this section, or elsewhere, does Colbath et al. disclose forming labels for the clusters from cluster lists that include topics that are associated with at least half of the documents in the clusters, as required by claim 1.

The Examiner also alleged that "as an identification system to cluster and give a unique name" (Office Action, page 4). Applicants cannot understand what the Examiner is alleging. If the Examiner persists with a rejection of the above-identified claim feature based on Colbath et al., Applicants respectfully request that the Examiner provide a reasonable explanation why the Examiner believes that the above-identified sections of Colbath et al. disclose forming labels for the clusters from cluster lists that include topics that are associated with at least half of the documents in the clusters, as required by claim 1.

For at least these reasons, Applicants submit that claim 1 is not anticipated by Colbath et al. Claims 2 and 3 depend from claim 1 and are, therefore, not anticipated by Colbath et al. for at least the reasons given with regard to claim 1.

Independent claims 9 and 16 recite features similar to (but possibly different in scope from) features recited in claim 1. Claims 9 and 16 are, therefore, not anticipated by Colbath et al. for at least reasons similar to reasons given with regard to claim 1.

*REJECTION UNDER 35 U.S.C. § 103(a)*

In paragraph 11 of the Office Action, the Examiner rejected claims 4-8, 10, 11, and 17 under 35 U.S.C. § 103(a) as allegedly unpatentable over Colbath et al. in view of Liddy et al. Applicants respectfully traverse the rejection.

Claims 4-8 depend from claim 1, claims 10 and 11 depend from claim 9, and claim 17 depends from claim 16. Without acquiescing in the Examiner's rejections with regard to claims 4-8, 10, 11, and 17, Applicants submit that the disclosure of Liddy et al. does not cure the deficiencies in the disclosure of Colbath et al. identified above with regard to claims 1, 9, and 16. For example, Liddy et al. does not disclose or suggest determining whether identified topics are associated with at least half of the documents in the clusters, or forming labels for the clusters from cluster lists that include topics that are associated with at least half of the documents in the clusters, as required by claim 1 (similar features are recited in claims 9 and 16). Instead, Liddy et al. discloses using headlines from newspaper articles or titles from documents in clusters to form labels for the clusters (column 25, lines 34-35).

Claims 4-8, 10, 11, and 17 are, therefore, patentable over Colbath et al. and Liddy et al., whether taken alone or in any reasonable combination, for at least the reasons given with regard to

claims 1, 9, and 16. Claims 4-8, 10, 11, and 17 are also patentable over Colbath et al. and Liddy et al. for reasons of their own.

For example, claim 4 recites assigning ranks to the ones of the topics based on a number of the documents with which the ones of the topics are associated. Neither Colbath et al. nor Liddy et al. discloses the combination of features of claim 4.

The Examiner admitted that Colbath et al. does not disclose assigning ranks, but alleged that Liddy et al. discloses assigning ranks and cited column 21, lines 28-52, of Liddy et al. for support (Office Action, page 6). Applicants disagree.

At column 21, lines 28-52, Liddy et al. discloses:

Matcher 55 matches documents by comparing the documents with the query and assigning each document a similarity score for the particular query. Documents with sufficiently high scores are arranged in ranked order in three folders, according to their relative relevance to the substance of a query. There are a number of evidence sources used for determining the similarity of documents to a query request, including:

Complex Nominals (CNs)\*

Proper Nouns (PNs)\*

Subject Field Codes (SFCs)

Single Terms\*

Text Structure

Presence of Negation

Mandatory requirements

\*CNs, PNs, and Single Terms are collectively called "terms."

Documents are arranged for the user based on a two-tier ranking system. The highest-level ranking mechanism is a system of folders. Documents are placed within folders based on various criteria, such as the presence or absence of mandatory terms. The lower-level ranking mechanism sorts documents within each folder based on criteria such as similarity score, document date assignment, etc.



In this section, Liddy et al. discloses ranking search result documents based on a query match.

Nowhere in this section, or elsewhere, does Liddy et al. disclose or suggest assigning ranks to topics, let alone assigning ranks to the ones of the topics based on a number of the documents with which the ones of the topics are associated, as required by claim 4.

For at least these additional reasons, Applicants submit that claim 4 is patentable over Colbath et al. and Liddy et al.

Claim 5 recites ranking the ones of the topics based on a number of the documents with which the ones of the topics are associated. Neither Colbath et al. nor Liddy et al. discloses the combination of features of claim 5.

The Examiner alleged that both Colbath et al. and Liddy et al. disclose these features and cited column 6, lines 42-49, of Colbath et al., and column 24, line 56 - column 25, line 2, of Liddy et al. for support (Office Action, page 7). Applicants disagree.

At column 6, lines 42-52, Colbath et al. discloses:

The OnTopic topic indexing system is a brand new, completely statistical system for topic classification. Previous work in the area of topic classification typically considered only one topic per segment, and modeled only a small number of topics. OnTopic has an HMM-based classification system, estimated from labeled training data. The topic HMM includes a model for every topic encountered in training and a model for general language that acts as an absorber of words that are not strongly associated with any topic.

In this section, Colbath et al. discloses an HMM-based classification system to identify topics in documents. Nowhere in this section does Colbath et al. even mention ranking topics, let alone ranking the ones of the topics based on a number of the documents with which the ones of the topics are associated, as required by claim 5.

At column 24, line 56 - column 25, line 2, Liddy et al. discloses:

The matching of documents to a query organizes documents by matching scores in a ranked list. The total number of presented documents can be selected by the user, the system can determine a number using the Recall Predictor (RP) function, or, in the absence of user input, the system will retrieve all documents with a non-zero score. Note that documents from different sources are interfiled and ranked in a single list.

The RP filtering function is accomplished by means of a multiple regression formula that successfully predicts cut-off criteria on a ranked list of relevant documents for individual queries based on the similarity of documents to queries as indicated by the vector matching (and optionally the proper noun matching) scores.

In this section, Liddy et al. discloses matching documents to a search query and organizing the documents by matching scores in a ranked list. Nowhere in this section, or elsewhere, does Liddy et al. even mention ranking topics, let alone ranking the ones of the topics based on a number of the documents with which the ones of the topics are associated, as required by claim 5.

For at least these additional reasons, Applicants submit that claim 5 is patentable over Colbath et al. and Liddy et al. Claims 6-8 depend from claim 5 and are, therefore, patentable over Colbath et al. and Liddy et al., whether taken alone or in any reasonable combination, for at least the reasons given with regard to claim 5.

*REJECTION UNDER 35 U.S.C. § 102(b) OR UNDER 35 U.S.C. § 103(a)*

In paragraph 12 of the Office Action, the Examiner rejected claims 18 and 19 under 35 U.S.C. § 102(b) as allegedly anticipated by, or in the alternative, as unpatentable over Colbath et al. Applicants respectfully traverse the rejection.

Amended independent claim 18 is directed to a method for creating labels for clusters of documents. The method comprises identifying topics associated with the documents in the clusters; determining whether the topics are associated with at least a predetermined number of the documents in the clusters; and generating labels for the clusters using ones of the topics that are associated with the at least a predetermined number of the documents in the clusters.

Colbath et al. does not disclose the combination of features recited in claim 18. For example, Colbath et al. does not disclose determining whether identified topics are associated with at least a predetermined number of the documents in the clusters. The Examiner alleged that Colbath et al. discloses this feature and cited column 5, lines 34-48, and column 8, lines 11-16, of Colbath et al. for support (Office Action, page 8). Applicants disagree.

Column 5, lines 34-48, of Colbath et al. has been reproduced above. In this section, Colbath et al. discloses identifying changes in speakers within an audio signal. Nowhere in this section, or elsewhere, does Colbath et al. disclose determining whether identified topics are associated with at least a predetermined number of the documents in the clusters, as required by claim 18.

Column 8, lines 11-16, of Colbath et al. is reproduced above. In this section, Colbath et al. discloses an HMM-based model for document retrieval. Nowhere in this section, or elsewhere, does Colbath et al. disclose determining whether identified topics are associated with at least a predetermined number of the documents in the clusters, as required by claim 18.

The Examiner also alleged that "a response is half-correct when the label is correct" (Office Action, page 8). Applicants cannot understand what the Examiner is alleging. If the Examiner persists with a rejection of the above-identified claim feature based on Colbath et al., Applicants respectfully request that the Examiner provide a reasonable explanation why the Examiner believes that the above-identified sections of Colbath et al. disclose determining whether identified topics are associated with at least a predetermined number of the documents in the clusters, as required by claim 18.

Colbath et al. also does not disclose generating labels for the clusters using ones of the topics that are associated with the at least a predetermined number of the documents in the clusters, as

further recited in claim 18. Instead, Colbath et al. discloses forming a title for a document from all of the topics associated with that document (column 10, lines 13-20).

The Examiner alleged that Colbath et al. discloses generating labels for the clusters using ones of the topics that are associated with the at least a predetermined number of the documents in the clusters, and cited column 8, lines 11-16, column 14, lines 29-33, and column 15, lines 1-10, of Colbath et al. for support (Office Action, page 8). Applicants disagree.

Column 8, lines 11-16, of Colbath et al. is reproduced above. In this section, Colbath et al. discloses an HMM-based model for document retrieval. Nowhere in this section, or elsewhere, does Colbath et al. disclose generating labels for the clusters using ones of the topics that are associated with the at least a predetermined number of the documents in the clusters, as required by claim 18.

Column 14, lines 29 - column 15, line 10, of Colbath et al. is reproduced above. In this section, Colbath et al. discloses that annotators can annotate more speech data, additional names for the name spotter, or additional topics for the topic classifier. Nowhere in this section, or elsewhere, does Colbath et al. disclose generating labels for the clusters using ones of the topics that are associated with the at least a predetermined number of the documents in the clusters, as required by claim 18.

The Examiner also alleged that "as to add more topics for the topic classifier" (Office Action, page 8). Applicants cannot understand what the Examiner is alleging. Adding topics to a topic classifier cannot reasonably be equated to generating labels for the clusters using ones of the topics that are associated with the at least a predetermined number of the documents in the clusters, as required by claim 18. If the Examiner persists with a rejection of this claim feature based on Colbath et al., Applicants respectfully request that the Examiner provide a reasonable explanation

why the Examiner believes that the above-identified sections of Colbath et al. disclose generating labels for the clusters using ones of the topics that are associated with the at least a predetermined number of the documents in the clusters, as required by claim 18.

For at least these reasons, Applicants submit that claim 18 is not anticipated by and is patentable over Colbath et al. Claim 19 depends from claim 18 and, therefore, is not anticipated by and is patentable over Colbath et al. for at least the reasons given with regard to claim 18.

### *CONCLUSION*

In view of the foregoing amendments and remarks, Applicant respectfully requests reconsideration and allowance of pending claims 1-11 and 16-19.

As Applicants' remarks with respect to the Examiner's rejections overcome the rejections, Applicants' silence as to certain assertions by the Examiner in the Office Action or certain requirements that may be applicable to such rejections (e.g., whether a reference constitutes prior art, motivation to combine references, etc.) is not a concession by Applicants that such assertions are accurate or such requirements have been met, and Applicants reserve the right to dispute these assertions/requirements in the future.

If the Examiner believes that the application is not now in condition for allowance, Applicants respectfully request that the Examiner contact the undersigned to discuss any outstanding issues.

Application No. 10/685,479  
Amendment dated November 15, 2006  
Reply to Office Action of August 15, 2006

Docket No.: BBNT-P01-199

Applicants believe no fee is due with this response. However, if a fee is due, please charge our Deposit Account No. 18-1945, under Order No. BBNT-P01-199 from which the undersigned is authorized to draw.

Dated: November 15, 2006

Respectfully submitted,

By 

Edward A. Gordon, Esq.

Registration No.: 54,130

FISH & NEAVE IP GROUP, ROPES & GRAY  
LLP

One International Place  
Boston, Massachusetts 02110-2624  
(617) 951-7000  
(617) 951-7050 (Fax)  
Attorneys/Agents For Applicant